



## Large Scale Molecular Dynamics Simulations for the Derivation of Solvation Free Energies of Strongly Hydrophobic Molecules

Reema Mahajan, Dieter Kranzlmüller, Jens Volkert,  
Siegfried Höfinger

published in

*NIC Workshop 2006,*  
*From Computational Biophysics to Systems Biology,*  
Jan Meinke, Olav Zimmermann,  
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)  
John von Neumann Institute for Computing, Jülich,  
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,  
ISBN-13: 978-3-9810843-0-6, pp. 173-176 , 2006.

© 2006 by John von Neumann Institute for Computing

Permission to make digital or hard copies of portions of this work for personal or classroom use is granted provided that the copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise requires prior specific permission by the publisher mentioned above.

<http://www.fz-juelich.de/nic-series/volume34>

# Large Scale Molecular Dynamics Simulations for the Derivation of Solvation Free Energies of Strongly Hydrophobic Molecules

Reema Mahajan<sup>1</sup>, Dieter Kranzlmüller<sup>2</sup>, Jens Volkert<sup>2</sup>, and Siegfried Höfner<sup>3</sup>

<sup>1</sup> Indian Institute of Technology,  
Department of Chemical Engineering,  
Hauz Khas, New Delhi-16, India  
*E-mail: reema.mahajan@gmail.com*

<sup>2</sup> Johannes Kepler University of Linz,  
Institute of Graphics and Parallel Processing  
Altenberger Straße 69, A-4040, Linz, Austria  
*E-mail: {kranzlmuller,volkert}@gup.uni-linz.ac.at*

<sup>3</sup> Michigan Technological University,  
Department of Physics,  
1400 Townsend Drive, Houghton, MI 49931, USA  
*E-mail: shoefing@mtu.edu*

Liquid water is believed to be of crucial biological importance when it comes to solvating hydrophobic molecules. The so-called hydrophobic effect is a key-determinant in the protein folding problem and a deeper understanding of it would certainly be of great value to the biophysics community. The process of solvating hydrophobic molecules in water may be studied with the help of computer simulations that establish an atomistic picture of the liquid state of water. A study has already been carried out in the past using Molecular Dynamics simulations and a polarizable force field description for the water molecules. The focus of that study was on the temperature dependence of the free energy change observed for the onset and growth of a repulsive sphere that represented the hydrophobic species. The obtained data was in good agreement to predictions from the Lum Chandler Weeks theory on the unit volume entropy. However, only small-sized hydrophobic solutes could be studied and the critical domain where LCW theory claims a change in underlying physics could not be examined. Therefore an extension of this previous study is currently underway using again a high level polarizable force field and a largely extended water box of dimension 60x60x60 Å. The technical complexity of such an approach is discussed and an outlook is given on the various implications of such an investigation.

## 1 Introduction

About 50 years have passed by since Walter Kauzmann has published his extensive review on thermodynamic principles of protein unfolding<sup>1</sup>. One of his conclusions has been that the so-called “Hydrophobic Bonds” would constitute the major stabilizing factor for protein assembly and protein stability. It is however till today that the actual kernel of hydrophobicity must still be considered an unsolved problem. Among other theoretical attempts Lum Chandler Weeks (LCW) theory of hydrophobicity has become fashionable. LCW theory divides small length scale hydrophobicity from large length scale hydrophobicity formally (volume dependence versus surface area dependence) as well as in terms of underlying physics (entropy drop versus entropy gain). The switch between these two

domains is said to occur at dimensions of  $r=10 \text{ \AA}$  of spherical model solutes. Since an unbiased experimental verification of such a picture of hydrophobicity is extremely difficult it appears to be only natural to address the underlying physics behind LCW theory with computational approaches.

## 2 Methods

Suitable computational techniques for this kind of problem are free energy calculations, which have been introduced and methodically refined to directly study  $\Delta G$  trends for various physico-chemical processes<sup>2</sup>. The obvious physical process to study here was the  $\Delta G$  corresponding to the introduction, onset and growth of an artificial repulsive sphere located in the center of a simulation cell otherwise filled with water molecules. The spherical repulsive cavity would then represent a model hydrophobic solute and the associated free energy change was the energy necessary to embed this hydrophobic solute in water. Exactly the same type of  $\Delta G$  has also become known as the cavitation free energy<sup>3</sup>.

A previous study has already been carried out employing the high level polarizable AMOEBA force field<sup>4</sup> for computer simulation of liquid water at seven individual temperatures in the range of 277K to 370K for calculating the free energy cost of creation of small-sized cavities (work currently under review). Ewald summation had been applied within the Molecular Dynamics simulations (MD) that form the basis of the Free Energy Perturbation (FEP) calculations in the framework of the Overlapping Spheres Technique (OST)<sup>5</sup>. Our previous study could reconfirm the Chandler picture of the unit volume entropy  $\Delta S$  for small-sized hydrophobic solutes (from the temperature profiles of  $\Delta G$ ). The computational demand of this study was on the order of three CPU years (Itanium 1.4 GHz). Volumes of up to  $r \approx 3 \text{ \AA}$  could be introduced, hence the small cavity regime could be examined.

In extension to this previous study our present efforts focus on largely increased volumes of up to  $r=10 \text{ \AA}$  to probe the critical domain where the switch in physical exemplification is thought to occur. We therefore need to increase our water box substantially and start out with a trial system of dimension  $60 \text{ \AA} \times 60 \text{ \AA} \times 60 \text{ \AA}$ . The CPU requirements of such an approach must be tremendous. In this current work we estimate the required computational resource allocation of such a large scale attempt and try to find a suitable compromise in case the resource allocation turns out to be simply un-feasible.

## 3 Results and Conclusions

The estimated computational resource allocation for the case of a MD/FEP study based on AMOEBA water boxes of dimension  $60 \text{ \AA} \times 60 \text{ \AA} \times 60 \text{ \AA}$  is summarized in Table 1 (middle column). Compared to the previously required computing resources (see Table 1, left column) this is however far beyond any acceptable limit. We therefore shrank the size of the water boxes layer by layer and estimated resulting overall resource allocations for various constellations. A cubic arrangement of about 1331 water molecules (see Table 1, right column) was finally chosen for production runs because it represents a good compromise between intended problem size and acceptable resource allocation. Such large scale computations can be conveniently carried out on large GRID systems in reasonable time lines.

	Previously ( $B \leq 3 \text{ \AA}$ )	Estimate ( $B \approx 10 \text{ \AA}$ )	Estimate ( $B \approx 8 \text{ \AA}$ )
# Water Molecules	216	8000	1331
Box Size	$(18.6\text{\AA})^3$	$(60\text{\AA})^3$	$(34.2\text{\AA})^3$
Ewald, Cut-Off Radius	14 $\text{\AA}$	14 $\text{\AA}$	14 $\text{\AA}$
Time Step	1 fs	1 fs	1 fs
$\Delta B$ in OST	$\pm 25 \times 0.0032 \text{ \AA}$	$\pm 25 \times 0.0032 \text{ \AA}$	$\pm 25 \times 0.0032 \text{ \AA}$
# Intervals in OST per $\text{\AA}$	12.5	12.5	12.5
Maximum Allowed Perturbation	2 RT	2 RT	2 RT
Annealing Period	2 ps	2 ps	2 ps
# Temperatures Considered	7	3	3
Intended MD/FEP Period	100 ps	100 ps	100 ps
# Trajectories per Temperature	29	120	90
Estimated CPU Time per Trajectory (Itanium 2, 1.4 Ghz)	5 days	5 years	2 months
# Trajectories Total	203	360	270
Estimated Total CPU Time	2.78 years	1800 years	45 years

Table 1. Experienced and estimated resource allocation for large scale MD/FEP simulations on cavitation free energy calculations in AMOEBA water. Previously consumed resources are shown in the column to the left, the middle column gives an estimate for a huge water box aimed at studies of cavities of radii up to  $B \approx 10 \text{ \AA}$ , and the right column shows estimates for a largely increased water box for cavitation studies of radii up to  $B \approx 8 \text{ \AA}$ . This final column to the right represents a good compromise between intended study size and feasible computational throughput.

## Acknowledgments

This work was supported in part by the National Institutes of Health Grant GM62838.

## References

1. W. Kauzmann, *Some factors in the interpretation of protein denaturation*, Adv. Protein Chem. **14**, 1–63 (1959).

2. T. Simonson, G. Archontis, M. Karplus, *Protein-Ligand Recognition: Free Energy Simulations Come of Age*, *Acc. Chem. Res.* **35**, 430–437 (2002).
3. J. Tomasi, B. Mennucci, R. Cammi, *Quantum Mechanical Continuum Solvation Models*, *Chem. Rev.* **105**, 2999–3094 (2005).
4. P. Ren, J. W. Ponder, *Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation*, *J. Phys. Chem. B* **107**, 5933–5947 (2003).
5. P. M. Postma, H. J. C. Berendsen, J. R. Haak, *Thermodynamics of cavity formation in water*, *Farad. Symp. Chem. Soc.* **17**, 55–67 (1982).